Carry-over of attentional settings between distinct tasks: A transient effect independent of top-down contextual biases

Catherine Thompson[1], Alessia Pasquini[1], and Peter J Hills[2]

[1] School of Health and Society, University of Salford, Salford, UK

[2] Department of Psychology, Bournemouth University, Poole, UK

Corresponding Author:

Dr Catherine Thompson

Allerton Building

University of Salford

Salford

M5 4WT

c.thompson@salford.ac.uk

Highlights

- Attentional settings from an initial task persist to a second, unrelated task, affecting subsequent attention and spread of search
- Top-down biases present within scenes have no impact on the persistence of attentional set from an initial task
- The influence of a previously relevant attentional set on eye movements in a second task is a brief effect

Abstract

Top-down attentional settings can persist between two unrelated tasks, influencing visual attention and performance. This study investigated whether top-down contextual information in a second task could moderate this “attentional inertia” effect. Forty participants searched through letter strings arranged horizontally, vertically, or randomly and then made a judgement about road, nature, or fractal images. Eye movements were recorded to the picture search and findings showed greater horizontal search in the pictures following horizontal letter strings and narrower horizontal search following vertical letter strings, but only in the first 1000ms. This shows a brief persistence of attentional settings, consistent with past findings. Crucially, attentional inertia did not vary according to image type. This indicates that top-down contextual biases within a scene have limited impact on the persistence of previously relevant, but now irrelevant, attentional settings.

Keywords

Attentional inertia; top-down; attentional set; eye movements; contextual information

1. Introduction

To allow for effective task completion an individual's limited cognitive resources will be biased towards relevant information and locations and away from irrelevant information and locations. This is achieved via the adoption of a top-down attentional set (e.g., Egeth, Virzi, & Garbart, 1984; Folk, Remington, & Johnston, 1992; Leber & Egeth, 2006). The top-down set supports selective attention (Johnston & Dark, 1986; Schneider & Shiffrin, 1977; Theeuwes, 1993) by prioritising specific stimuli and areas of space on the basis of task demands. It therefore dictates which stimuli will be selected and which will be ignored. Findings show that these 'attentional control settings' (Folk et al., 1992) influence the allocation of attention and therefore performance in a task (e.g., Kaptein, Theeuwes, & van der Heijden, 1995; Rossi & Paradiso, 1995).

When a task changes the attentional settings should be updated to reflect this, however this is not always the case. Studies have shown that after responding to specific targets over a large number of trials, participants are unable to inhibit these targets when they subsequently become irrelevant (e.g., Leber & Egeth, 2006; Thompson, Underwood, & Crundall, 2007). This is argued to be due to the persistence of the top-down attentional set; the stimuli that were previously task-relevant continue to match the attentional settings and so resources are allocated to them. Leber and Egeth (2006) suggested that the failure to change set is due to the fact that the costs associated with switching set (greater investment of cognitive resources in order to inhibit the old settings) outweighs the benefit to performance of switching (in their study this benefit was the reduced contingent capture by irrelevant peripheral distracters). They supported this proposal by showing that an attentional set would persist from one task to a second task following 320 trials in the first task, but not following only 40 trials. The increased experience serves to consolidate the set, making it more difficult to update in line with new task demands, requiring more resources to inhibit.

The findings of Leber and Egeth (2006) suggested that perseveration of an attentional set would only occur following significant investment in the initial set. This would indicate that in more dynamic tasks, persistence of top-down settings would not pose a problem. However, newer findings show that this is not the case. For example, Wendt, Kahler, Luna-Rodriguez, and Jacobsen (2017) have demonstrated the persistence of spatial visual attention between two tasks of short duration. In an initial task participants were presented with three letters and were asked to identify the central target, therefore adopting a narrow focus of attention due to the need to inhibit flanking distracters, or they were asked to judge whether the three letters were the same or different, therefore adopting a wide focus of attention to allocate resources to all items. Following this first task participants were given a ‘probe’ task in which three digits were presented to the same locations as the letters in the preceding task and participants had to identify a particular target that was located in the centre or the periphery of the digit string. They found that central targets were identified faster than peripheral targets when the probe task was preceded by the flanker task, but the difference between response times to central and peripheral targets was much smaller following the same/different task. This shows that the spatial spread of attention adopted for one task can persist to a second task, even when this leads to a detriment to performance.

The study by Wendt et al. (2017) shows even limited exposure to an initial task can result in the persistence of attentional settings to a second task, yet their paradigm incorporated relatively simple stimuli presented to the same spatial location across all trials. Other work has adopted methods that are better able to reflect dynamic, real-world situations. Thompson and Crundall (2011) developed a set-switching paradigm in which participants searched through letter strings presented horizontally, vertically, and randomly across the screen and then searched a real-world image. To measure the impact of the letter strings on the allocation of attention to the images eye movements were recorded and results showed that spatial visual attention in this second task was influenced by the orientation of the preceding letter search. Vertical search in the picture task increased following a vertical letter search and decreased following a horizontal letter search.

These studies measure attentional set switching, yet this often occurs in tandem with task switching because when a task changes this requires an update in the internal rules used to respond to specific stimuli (an intentional set) and an update in the rules used to select specific stimuli for processing (an attentional set) (Rushworth, Passingham, & Nobre, 2002). In a standard task switching paradigm participants are shown stimuli and are asked to use one of two rules in order to respond to the stimuli. For example, when presented with a blue letter, participants respond to whether it is an uppercase or lowercase letter, when presented with a red letter, participants respond to whether it is a vowel or a consonant. In 'no-switch' trials the colour will be consistent (e.g. a red letter followed by another red letter), in 'switch' trials the colour will be different (e.g. a red letter followed by a blue letter). Numerous studies have shown that response times are longer on switch trials compared to no-switch trials and the difference between the two trial types is referred to as a switch cost (e.g., Meiran, 1996; Rogers & Monsell, 1995; Wylie & Allport, 2000). Allport, Styles, and Hsieh (1994) accounted for switch costs using the task-set inertia hypothesis which states that the old set will persist to a new task and the resources required to inhibit the set means there are fewer resources available for the new task.

It is notable that attentional set switching leads to performance switch costs in a similar way to task-switching. This is seen in the study by Wendt et al. (2017) but it has also been found in experiments utilising the paradigm developed by Thompson and Crundall (2011). For instance, when participants were asked to detect hazards in driving stimuli, not only did the orientation of the preceding letter search influence eye movements to the driving stimuli, but response times were slower following a vertical letter search compared to following a horizontal letter search (Hills, Thompson, & Pake, 2018; Thompson & Crundall, 2011). A vertical letter search has additionally been found to negatively impact the ability to recognise upright faces (Hills, Mileva, Thompson, & Pake, 2017). It may be argued that this impairment to performance reflects an effect of fatigue rather than attentional shifting, on the basis that a vertical search is more difficult, takes longer, and thus leads to performance costs in a subsequent task. However, if that was the case it would also be expected that carry-over would be greater with increased repetition of the letter search task, yet studies have consistently shown that the persistence of attentional set to the second task does not vary according to whether participants complete one or three letter searches (Hills et al., 2017; Hills et al., 2018; Thompson & Crundall, 2011).

The influence of persisting attentional settings on eye movements has also been found by Longman, Lavric, and Monsell (2013). They presented participants with faces that each contained a letter on the forehead. In each trial participants were cued to identify the letter or the face and their results showed that on switch trials participants were more likely to fixate the previously-relevant region. They referred to this effect as “attentional inertia”. Longman, Lavric, and Monsell (2017) propose that attentional inertia is a component of task-set inertia but the two can be separated; attentional inertia being the persistence of where/what the attentional resources should be allocated towards, and task-set inertia being the persistence of the rules for how to respond to the attended information. They used a gaze-contingent paradigm in which participants moved their eyes from a fixation cross to a stimulus that appeared in different areas of the screen. The stimulus only appeared once participants moved their eyes from fixation and they were more likely to move towards areas that previously contained target stimuli, leading to slower response times to targets appearing in a different location to that of the previous trial, compared to when target location was repeated across trials. However, they also found that when participants were given an unlimited amount of time to prepare for a trial (i.e. they remained fixated until they were ready) there was no evidence of attentional inertia and participants did not re-fixate previously relevant areas, yet the performance switch costs still remained. It is therefore worth investigating attentional inertia as an effect independent to task-set inertia.

Within the task switching literature there is reference to switching involving activation of the new set and disengagement from and inhibition of the old set (e.g. Rogers & Monsell, 1995). Studies showing that carry-over increases due to investment in the initial set (e.g., Leber & Egeth, 2006; Thompson, Howting, & Hills, 2015) reveal that difficulty in inhibiting an old set plays a specific role in the persistence of top-down settings, but so far it is unclear whether the activation of the new set has any influence on attentional inertia. There is some support for this influence, notably within the variation in findings regarding attentional inertia. In the studies completed by Thompson and Crundall (2011) the carry-over of search behaviour between the letter and picture search tasks only influenced vertical spread of search. Thompson et al. (2015) argued that this was because the images used in those experiments were road scenes which evoke a horizontal spread of search (e.g., Crundall & Underwood, 1998; Konstantopoulos, Chapman, & Crundall, 2010) and the biases already present in the images would interact with the carry-over from the letter search. Specifically, due to the familiar context of road images observers would quickly extract the scene ‘gist’ (Friedman, 1979; Oliva, 2005). This would then activate an attentional set suitable for such a scene (allocating resources towards the horizontal axis due to the statistical likelihood of relevant objects occurring in this area) and this set would moderate the impact of the attentional settings from the preceding task.

Scene gist is the basic information that allows a scene to be categorised (e.g., Oliva, 2015; Oliva & Torralba, 2001; Wu, Wang, & Pomplun, 2014), it can be extracted in less than 100ms, and it has a key influence on the guidance of attention and eye movements in natural scenes (e.g. Torralba, Oliva, Castelhano, & Henderson, 2006). There is substantial evidence that the semantic context of a scene can bias attention (and visual search). Foulsham, Kingstone, and Underwood (2008) demonstrated that in natural, outdoor scenes participants had a tendency to make more horizontal eye movements and this bias towards the horizontal axis shifted as the images were rotated. This contrasted with indoor scenes for which there was little difference between the proportion of horizontal and vertical eye movements. These findings support the notion that different stimuli incorporate different top-down information, and this biases attention to certain locations and objects. This aligns with work showing that, on the basis of past experience, attention is prioritised to areas that are more likely to contain task-relevant information (e.g., Brockmole & Henderson, 2006; Hills, Sullivan, & Pake, 2012; Shinoda, Hayhoe, & Shrivastava, 2001).

The main aim of the present work was to test the influence of existing top-down semantic (contextual) signals on the activation of an attentional set. There is now a body of evidence showing that top-down attentional settings can persist from a task in which they are relevant to a task in which they are no longer relevant, and it has been argued that attentional set switching should be explored separately to task switching. Yet, unlike task switching, the mechanisms involved in attentional set switching are not fully understood, beyond knowing that switching is more effortful when more resources have been invested in the initial set (i.e., Leber & Egeth, 2006; Thompson, et al., 2015) and that preparation can support set switching (Longman et al., 2017). Such findings emphasise the role of inhibition in successful set switching. Building on past work and taking advantage of findings showing that attention varies due to contextual information, the present study will explore whether, like task switching, attentional set switching is impacted by the activation of a new set, in addition to the inhibition of an old set. This will show the relative importance of activation and inhibition within attentional set switching and will also indicate whether some contexts are more at risk from attentional inertia than others.

The paradigm of Thompson and Crundall (2011) was utilised but the type of images used in the second task was manipulated. In every trial participants were presented with nine letters arranged horizontally, vertically, or randomly across the screen and were asked to search the letters to determine if there were three or four vowels presented. They completed either a single letter search or three letter searches (different letters but presented in the same ‘orientation’ within each trial) before being shown an image for 4000ms. They were asked to view this image to make a judgement about how complex they found it and the images shown were road scenes, nature scenes, or fractal images. The use of fractals followed the work of Foulsham and Kingstone (2010) who found that participants made more horizontal eye movements on fractal images and made more vertical eye movements on nature images. However, whilst search in nature scenes varied according to the angle with which the scene was rotated, eye movements to fractals remained constant across all angles of rotation. This indicates that fractals do not contain semantic information to guide attention and so they are free of top-down influences (the consistent biasing of attention to the horizontal axis is instead representative of oculomotor behavioural biases, e.g. Tatler & Vincent, 2009). Given the argument that attentional set switching involves activation of the new set (in addition to inhibition of the old set) it was predicted that the attentional inertia effect would be smaller for images that contain more top-down information (i.e. road images) compared to images that contain no semantic information (fractals). This is because the contextual guidance in the road images would allow for quick activation of a new attentional set overriding any persisting influence from the previous task. The lack of semantic context in fractal images will mean that an attentional set is not activated, allowing for greater carry-over from the letter search task.

An additional aim of the current work was to further explore the duration of the carry-over effect by comparing eye movements in the first 1000ms of the picture search task with those in the second 1000ms. Thompson et al. (2015) found that whilst orientation of letter strings influenced initial eye movements to a picture search, this carry-over was short-lived and did not last beyond 1000ms. The present experiment will provide further evidence for the time course of attentional inertia and will also demonstrate whether the duration of the effect is impacted by the top-down information in the second task. Three measures of carry-over were used, horizontal and vertical spread of search and saccade direction. Spread of search was calculated as the standard deviation of the horizontal and vertical positions of each fixation and saccade direction was the percentage of saccades made in an upwards, downwards, leftwards, or rightwards direction. The measures were compared for each level of orientation of the letter search and for the first 1000ms and the second 1000ms in the picture search. In line with previous findings, it was predicted that attentional inertia would be more apparent in the first 1000ms of the picture search task. However, the effect would last beyond 1000ms when participants are searching the fractal images due to the lack of contextual information to guide attention and override the previously relevant settings.

2. Material and methods

2.1 Design

The experiment was completed using a 3x3x2 within-participants design. The first independent variable was *orientation* of the stimuli presented in the letter search task (horizontal, vertical, and random). The second independent variable was the *image type* presented after the letter search (road scenes, nature scenes, and fractal images). The third independent variable was *time*. Eye movements in the picture search task were separated into two time epochs of the first 1000ms and the second 1000ms. The dependent variables consisted of horizontal spread of search, vertical spread of search, and saccade direction. Horizontal and vertical spread of search were calculated using the standard deviation of the "x" and "y" coordinates of each fixation (measured in degrees). To measure saccade direction the angle of each saccade was calculated and then each saccade was coded into one of four bins: upwards (315°-45°), rightwards (45°-135°), downwards (135°-225°), and leftwards (225°-315°). A mean complexity rating was taken for each image type for each participant, ranging from 1 (very simple) to 6 (very complex).

2.2 Participants

Participants were an opportunity sample of 40 students (24 female) from the University of Salford who completed the experiment for an inconvenience allowance of £5. Participants were aged between 19 and 42 and the mean age was 29.05 (SD = 6.90). All participants reported normal or corrected-to-normal vision.

2.3 Stimuli and Apparatus

The experiment was designed and run on a Viglen genie Intel Core Duo 2 computer with a 22-inch screen using E-Prime Pro 2. A chin rest was used and participants were seated 60cm from the screen. A Tobii X50 (Stockholm, Sweden) eye-tracker recorded eye movements. This had a sampling rate of 50Hz, the minimum fixation duration was 100ms, and the minimum fixation dispersion threshold was 100 pixels. In the task participants were presented with letters in black on a white background in Verdana font size 18, subtending

0.95° x 0.95°. Nine letters were shown in each display, with 3 vowels and 6 consonants, or 4 vowels and 5 consonants. The letters could be upper or lowercase and all letters were used with the exception of the letter I. An invisible 9x9 grid was used to locate the letters in each display (28° x 28°), with stimuli presented across the centre of the display in a horizontal search, down the centre of the screen in a vertical search, and arranged randomly in a random search. A total of 144 images were presented, 48 road scenes, 48 nature scenes, and 48 fractal images. All images measured 35.14° x 28.07° and were shown in colour. The road and nature scenes were taken in and around Greater Manchester, UK and the fractal images were obtained from the Spanky Fractal Database. See figure 1 for an example of the images used within the carry-over task.

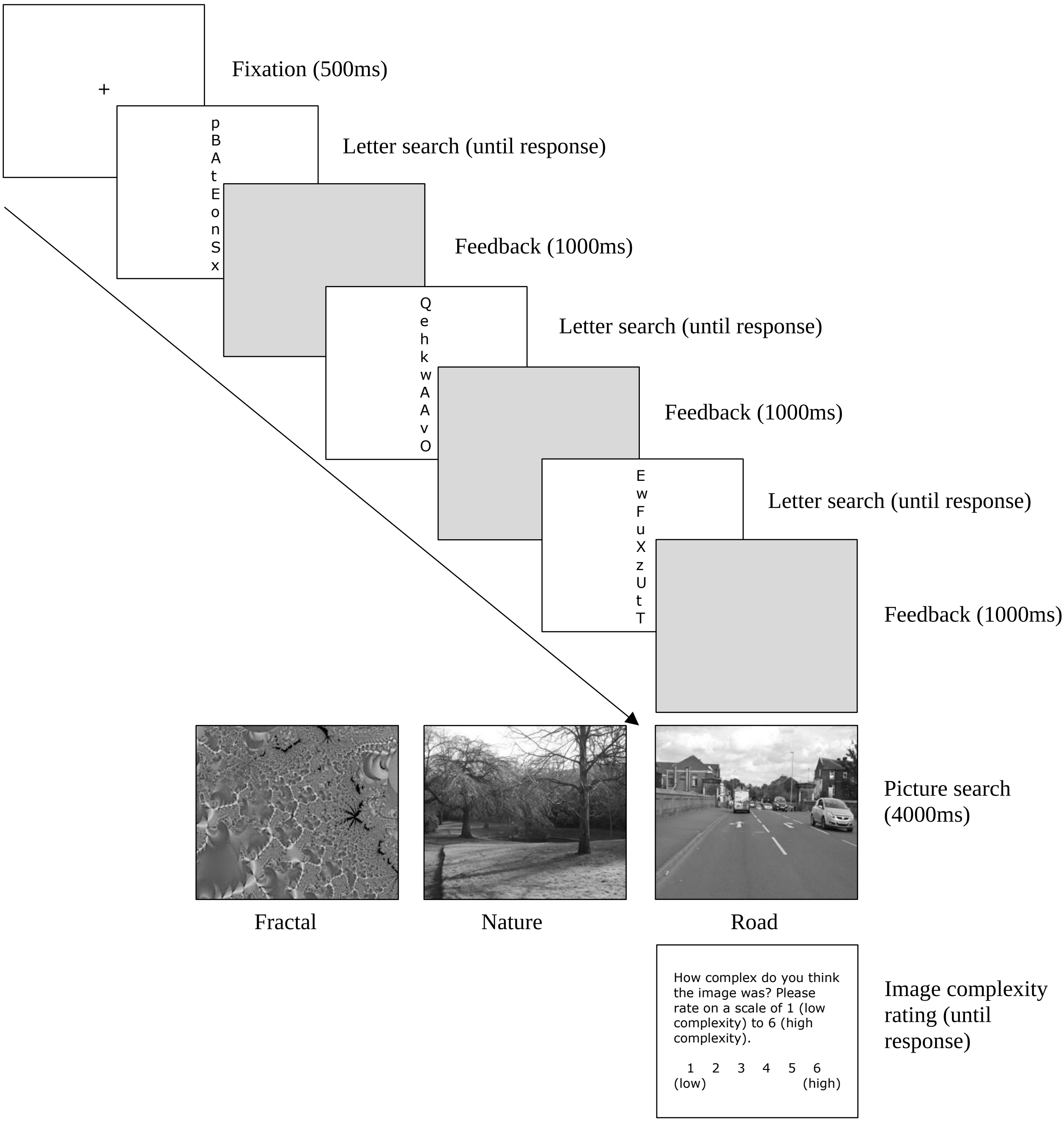

Figure 1: The sequence of events in a single trial showing the letter search task followed by the picture search task. A single picture (either a fractal, nature, or road image) was presented after one or three repetitions of the letter search and the example here shows a vertical letter search with three repetitions. Participants were asked to rate the image for complexity at the end of each trial. All images were shown in full colour.

2.4 Procedure

When arriving for the experiment participants were asked to sit with their chin in the chin rest and were then given on-screen instructions for the task. Their eye movements were then calibrated using a 5-point calibration screen. In each trial (see figure 1) a fixation cross appeared at the centre of the screen for 500ms. A letter search was then presented with letters arranged horizontally, vertically, or randomly across the screen. Participants were asked to search through the letter string and respond to the number of vowels by pressing the keys “3” and “4” on the keyboard with their left or right index finger. They were asked to respond as quickly as possible and the letters remained on the screen until a response was made. Each response was followed by a blank feedback screen for 1000ms (green for correct responses and red for incorrect responses). In half the trials two further letter searches were completed with letters in the same orientation before a picture was presented. In the other half a picture was presented after a single letter search[1]. Participants were asked to view each picture in a natural way in order to make a response regarding its complexity. They were told that complexity was the level of intricate detail in the picture and after viewing an image for 4000ms they were given a rating scale and asked to press a key from 1 (low complexity) to 6 (high complexity) to indicate their response. Participants completed 144 trials in this task. There were 48 trials for each orientation of the letter search task with 16 trials for each image type. In 8 of these trials a single letter search was presented and in 8 trials three letter searches were presented. There were an equal number of letter searches with 3 and 4 vowels for each trial-type. Trials were presented in a random order.

---

[1] Thompson et al. (2015) found that 4 and 8 repetitions of a letter search prior to a picture increased carry-over (compared to 1 repetition), however Thompson et al. (2011) found no effect of repetition when using 1, 2, and 3 letter displays prior to a picture. Consistent with Thompson et al. (2011), preliminary analysis showed no effect of repetition on carry-over, therefore it was not included as an independent variable and the analysis is not presented here.

3. Results

Analysis of the carry-over task was conducted on eye-movements to the picture search. Eye movements were recorded for the full 4000ms that each image was presented, however analysis was restricted to eye-movements made in the first 2000ms and these eye movements were separated into two epochs (first 1000ms and second 1000ms) to allow an investigation of the duration of the carry-over effect. Consistent with past research, the first full fixation and saccade made to the picture in each trial was removed. This was to ensure that the carry-over measures were not simply reflecting the eyes returning to the centre of the screen (e.g. following a vertical search where a participant makes a final fixation on the last, bottom-most letter, they may saccade back to the centre of the screen in preparation for the next task and this would reflect the procedure of the experiment, rather than any carry-over). Accuracy was recorded for the letter search and all trials in which the search immediately preceding a picture was incorrect were removed (8.33% of trials). This did not vary according to orientation, image type, or the number of repetitions of the letter search (all ps > .05).

Prior to analysing the eye movements made to the picture search the complexity ratings given to each image type were analysed using a within-participants ANOVA followed by simple planned contrasts to compare complexity of road and nature scenes to complexity of fractal images. The effect of image type was significant, F (2,78) = 86.362, MSE = 0.652, p < .001, *partial* η^2 = .689. Participants rated the fractal images (M = 4.76) as significantly more complex than the road scenes (M = 2.98; F (1,39) = 83.685, MSE = 1.509, p < .001, *partial* η^2 = .682) and significantly more complex than the nature scenes (M =2.88; F (1,39) = 141.634, MSE = 1.013, p < .001, *partial* η^2 = .784).

Three dependent variables were selected for eye movements made to the picture search, spread of search along the horizontal axis, spread of search along the vertical axis, and saccade direction. Spread of search was analysed using two 3 (orientation) x 3 (image-type) x 2 (time) within-participants ANOVAs. These were followed by simple planned contrasts comparing search following a horizontal and vertical letter string to that following a random letter string and comparing search in road and nature images to search in fractal images. Where sphericity was violated the Greenhouse-Geisser values are reported but uncorrected degrees of freedom are presented unless this altered the level of significance.

For horizontal spread of search there was a significant effect of orientation (see figure 2), F (2,78) = 12.077, MSE = 0.420, p < .001, *partial* η^2 = .236. Contrasts showed significantly wider horizontal search on the pictures following a horizontal letter string (M = 3.15°) compared to a random letter string (M = 2.94°), F (1,39) = 10.721, MSE = 0.346, p = .002, *partial* η^2 = .216. Horizontal search on the pictures did not differ significantly following a vertical (M = 2.88°) or random letter string, F (1,39) = 1.274, MSE = 0.238, p = .266, *partial* η^2 = .032. There was a significant effect of image type, F (2,78) = 17.596, MSE = 1.211, p < .001, *partial* η^2 = .311. The spread of search along the horizontal plane was significantly wider in road images (M = 3.32°) than fractal images (M = 2.91°), F (1,39) = 15.652, MSE = 0.854, p < .001, *partial* η^2 = .286, but horizontal search did not differ between nature images (M = 2.74°) and fractals, F (1,39) = 2.725, MSE = 0.863, p = .107, *partial* η^2 = .065. There was also a significant effect of time, F (1,39) = 255.791, MSE = 1.036, p < .001, *partial* η^2 = .868. Horizontal spread of search was narrower in the first 1000ms of the picture search (M = 2.38°) and wider in the second 1000ms (M = 3.6°). Crucially there was no interaction between orientation and image type, F (4,156) = 0.314, MSE = 0.477, p = .869, *partial* η^2 = .008. Given the prediction of a smaller carry-over effect for images containing more top-down contextual information (i.e. road images) compared to images that contain no semantic information (fractals), this non-significant effect was also examined by estimating a Bayes factor using Bayesian Information Criteria (Wagenmakers, 2007). This compared the fit of the data under the null hypothesis (there would be no significant interaction between orientation and image type) and the alternative hypothesis (the effect of orientation would differ due to image type). Using JASP (JASP Team, 2020), adopting the default priors and comparing to the best model, the estimated Bayes factor was calculated as BF_{01} = 1.320e+76. This offers decisive support for the null hypothesis.

In relation to the duration of the carry-over, there was a significant interaction between orientation and time for horizontal spread of search, $F\ (2,78) = 4.072$, $MSE = 0.365$, $p = .021$, *partial* $\eta^2 = .095$. The planned contrasts showed that the relative difference in horizontal search following horizontal and random letter strings was consistent across epochs 1 and 2, $F\ (1,39) = 1.026$, $MSE = 0.255$, $p = .317$, *partial* $\eta^2 = .026$, with increased horizontal search after horizontal letter strings. However, although there was no difference between horizontal search following vertical and random letters across the first 2000ms of picture viewing, when this was broken down into separate epochs there was a significant difference, $F\ (1,39) = 7.816$, $MSE = 0.248$, $p = .008$, *partial* $\eta^2 = .167$. This showed reduced horizontal search following a vertical letter string in the first 1000ms of picture viewing (compared to following a random letter string), but no difference in the second 1000ms.

There was also an interaction between image type and time, showing that the differences between horizontal search across the images was more pronounced in the second 1000ms of viewing, $F\ (2,78) = 11.436$, $MSE = 0.678$, $p < .001$, *partial* $\eta^2 = .227$. Horizontal search was wider in road images than fractal images, but only in the second epoch, $F\ (1,39) = 15.848$, $MSE = 0.611$, $p < .001$, *partial* $\eta^2 = .289$. There was no interaction between orientation, image type, and time, $F\ (4,156) = 0.760$, $MSE = 0.399$, $p = .553$, *partial* $\eta^2 = .019$.

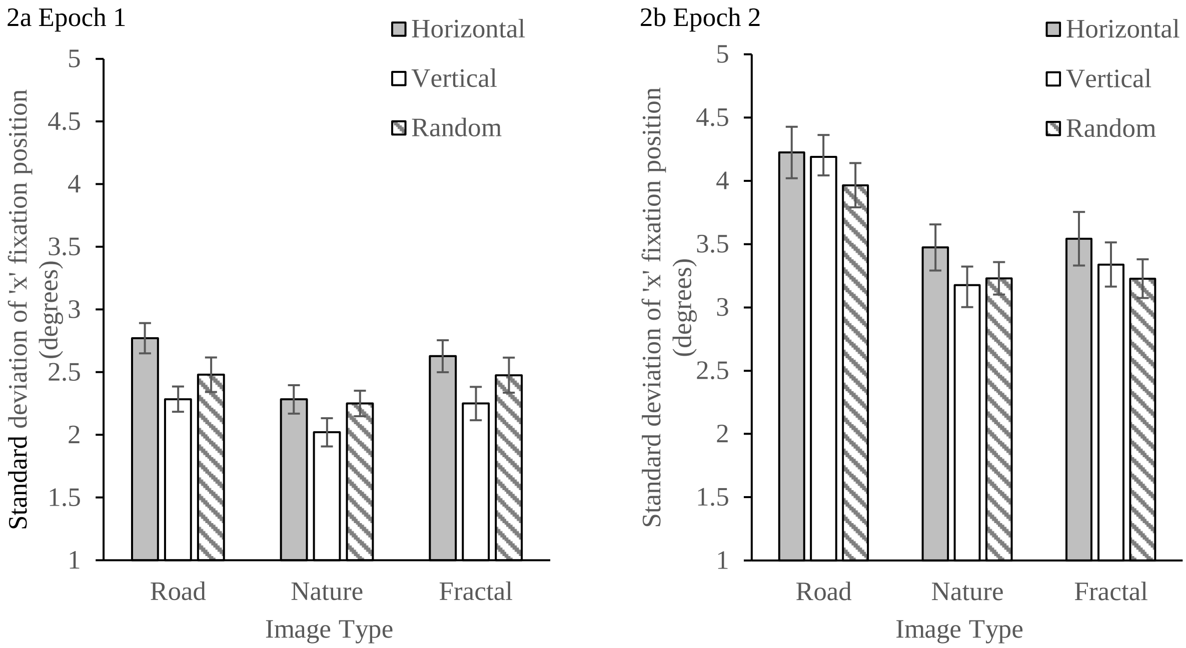

Figure 2: Spread of search along the horizontal axis in epochs 1 (a) and 2 (b) of the picture search based on the orientation of the preceding letter search task (horizontal, vertical, or random). Error bars represent standard error (calculated as the standard deviation/√40).

For vertical spread of search (see figure 3) there was no significant main effect of orientation, F (2,78) = 1.230, MSE = 0.176, p = .298, *partial* η^2 = .031. There was however a significant effect of image type, F (2,78) = 76.972, MSE = 0.645, p < .001, *partial* η^2 = .664. Vertical search was significantly narrower on a road image (M = 1.29°) compared to a fractal image (M = 2.05°), F (1,39) = 66.381, MSE = 0.694, p < .001, *partial* η^2 = .630. There was no significant difference between vertical spread of search to nature images (M = 2.11°) and fractals, F (1,39) = 1.027, MSE = 0.230, p = .317, *partial* η^2 = .026. Again, there was a significant effect of time, F (1,39) = 48.854, MSE = 0.432, p < .001, *partial* η^2 = .556, with greater vertical search in the second epoch (M = 1.99°) compared to the first (M = 1.65°). Consistent with horizontal spread of search, there was no interaction between orientation and image type, F (4,156) = 0.192, MSE = 0.238, p = .943, *partial* η^2 = .005. Again, the data contributing to this non-significant effect were also examined by estimating a Bayes factor (BF_{01} = 6.276e+20). This provided strong support in favour of the null hypothesis.

Whilst there was no main effect of orientation on vertical spread of search, there was a significant interaction between orientation and time, F (2,78) = 5.515, MSE = 0.151, p = .006, *partial* η^2 = .0124. In epoch 1 vertical search was significantly narrower following horizontal letter strings (M = 1.55°) compared to random letter strings (M = 1.71°), whereas in epoch 2 there was no difference (means of 2.02° and 1.98° respectively), F (1,39) = 7.564, MSE = 0.108, p = .009, *partial* η^2 = .162. There was no difference between vertical search following vertical and random letter strings across both epochs, F (1,39) = 0.004, MSE = 0.099, p = .952, *partial* η^2 = .000.

Similar to horizontal search, there was a significant interaction between image type and time for vertical spread of search, F (2,78) = 14.387, MSE = 0.321, p < .001, *partial* η^2 = .269. In this case, the significantly narrower vertical search on road images compared to fractals was much more pronounced in epoch 2 (means of 1.32 and 2.35) than in epoch 1 (means of 1.55 and 1.71), F (1,39) = 25.677, MSE = 0.232, p < .001, *partial* η^2 = .397. There was no interaction between orientation, image type, and time, F (4,156) = 0.492, MSE = 0.164, p = .741, *partial* η^2 = .012.

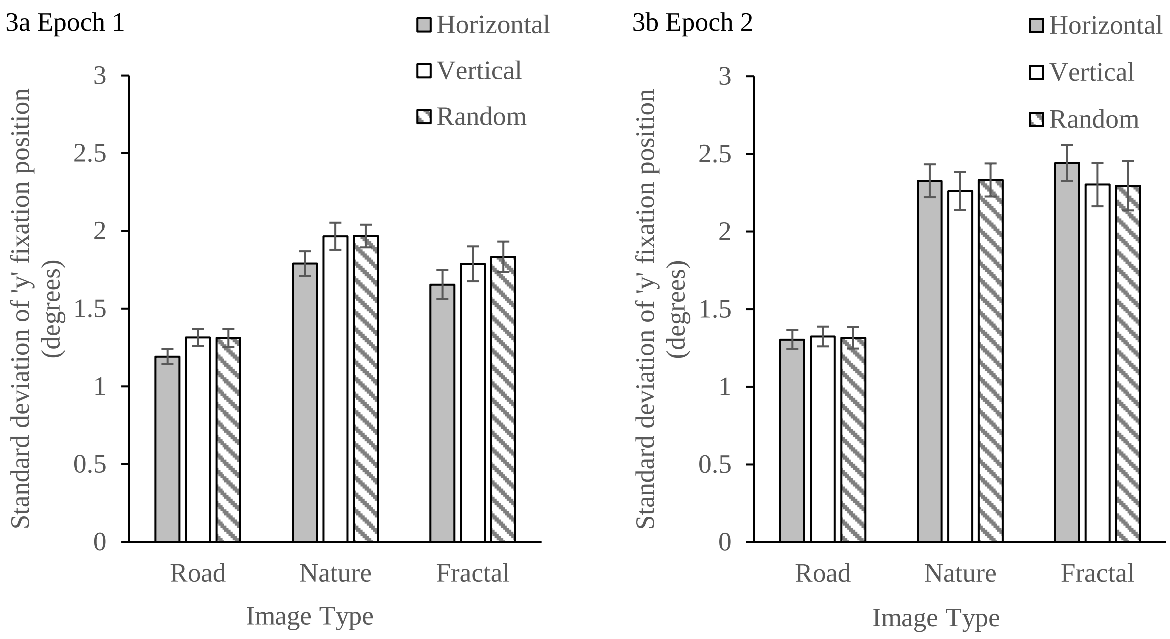

Figure 3: Spread of search along the vertical axis in epochs 1 (a) and 2 (b) of the picture search based on the orientation of the preceding letter search task (horizontal, vertical, or random). Error bars represent standard error (calculated as the standard deviation/√40).

To analyse the effect of the letter search task on saccade direction saccades in the first 2000ms of the picture search were coded into one of 4 bins based on direction (whether the saccade was made in an upwards, downwards, leftwards, or rightwards direction). The percentage of saccades in each bin was then calculated for each condition of orientation. This was analysed using a 4 (direction) x 3 (orientation) x 3 (image type) x 2 (time) ANOVA followed by the same planned contrasts for orientation and image type as those used for the spread of search analysis. Any significant effects of direction were further explored using deviation contrasts that compared the percentage of saccades in bins 1, 2, and 3 (upwards, rightwards, and downwards) to the mean (bin 4 was not included to control for multiple comparisons).

There was a significant effect of direction, F (3,117) = 259.640, MSE = 241.552, p < .001, *partial* η^2 = .869. Compared to the mean, participants made significantly fewer upwards saccades (*M* = 17.37%), F (1,39) = 214.557, *MSE* = 21.737, $p < .001$, *partial* $\eta^2 = .846$, significantly more rightwards saccades (*M* = 33.37%), F (1,39) = 231.360, *MSE* = 24.239, $p < .001$, *partial* $\eta^2 = .856$, and significantly fewer downwards saccades (*M* = 16.49%), F (1,39) = 370.107, *MSE* = 15.660, $p < .001$, *partial* $\eta^2 = .905$.

There was no interaction between direction and orientation, F (6,234) = 1.627, *MSE* = 78.335, $p = .140$, *partial* $\eta^2 = .040$, however there was a significant interaction between direction, orientation, and time, F (6,234) = 2.981, *MSE* = 105.051, $p = .008$, *partial* $\eta^2 = .071$. To analyse this further two separate ANOVAs were completed, one for saccades in the first 1000ms (figure 4a) and one for those in the second 1000ms (figure 4b).

Just taking eye movements in the first epoch, the interaction between direction and orientation was significant, F (6,234) = 2.607, *MSE* = 167.741, $p = .031$, *partial* $\eta^2 = .063$. This showed fewer rightwards saccades following a vertical letter search (34.53%) compared to following a random letter search (37.34%), F (1,39) = 4.405, *MSE* = 71.800, $p = .042$, *partial* $\eta^2 = .101$, and fewer downwards saccades following a horizontal letter search (15.64%) compared to following a random letter search (18.7%), F (1,39) = 11.527, *MSE* = 32.338, $p = .002$, *partial* $\eta^2 = .228$. In the second epoch the interaction between direction and orientation was non-significant, F (6,234) = 1.939, *MSE* = 55.968, $p = .075$, *partial* $\eta^2 = .047$.

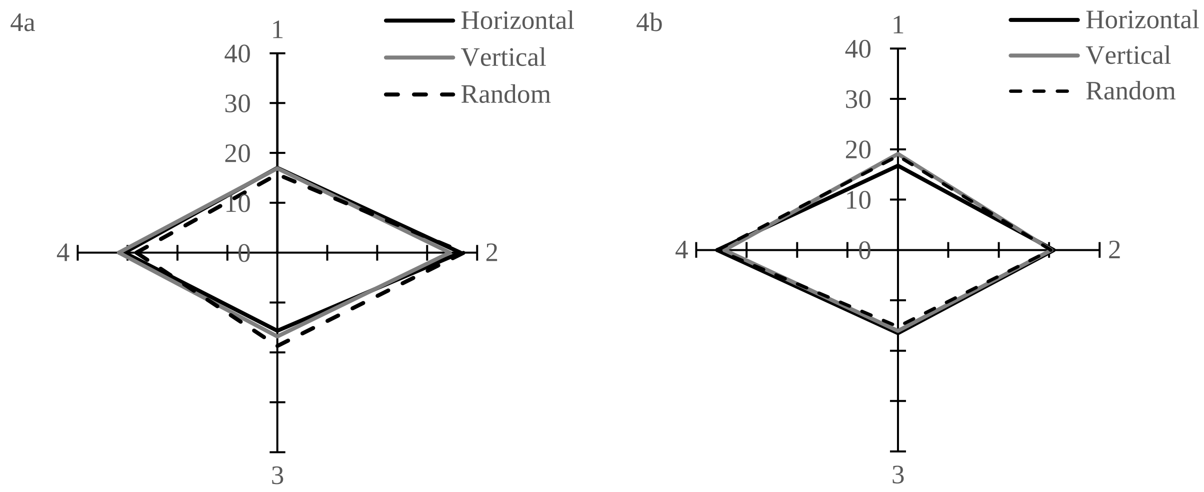

Figure 4: Percentage of saccades made in each of the four directions in (a) the first 1000ms of the picture task and (b) the second 1000ms of the picture task, across the three conditions of letter search orientation. Bins 1, 2, 3, and 4 identified on the axis represent upwards, rightwards, downwards, and leftwards saccades respectively.

Across the first 2000ms of the picture search there was a significant interaction between image type and direction (see figure 5), F (6,234) = 67.120, MSE = 128.385, p < .001, *partial* η^2 = .632. Compared to fractal images, on the road images participants made significantly fewer upwards saccades, F (1,39) = 41.990, MSE = 75.615, p < .001, *partial* η^2 = .518, significantly more rightwards saccades, F (1,39) = 94.163, MSE = 88.380, p < .001, *partial* η^2 = .707, and significantly fewer downwards saccades, F (1,39) = 145.488, MSE = 63.958, p < .001, *partial* η^2 = .789. Participants also made significantly more upwards saccades on nature images compared to fractals, F (1,39) = 14.938, MSE = 40.500, p < .001, *partial* η^2 = .277, and significantly fewer downwards saccades on nature images compared to fractals, F (1,39) = 6.485, MSE = 43.865, p = .015, *partial* η^2 = .143. This pattern of eye movements did not differ between epochs 1 (figure 5a) and 2 (figure 5b), F (6,234) = 1.503, MSE = 148.950, p = .178, *partial* η^2 = .037. Finally, there was no interaction between orientation, image type, and direction, F (12,468) = 1.175, MSE = 84.329, p = .298, *partial* η^2 = .029. Consistent with the spread of search analysis, the estimated Bayes factor used to examine this non-significant effect provided strong support for the null hypothesis compared to the alternative hypothesis (BF_{01} = 59840.472). There was also no interaction between orientation, image type, direction, and time, F (12,468) = 1.082, MSE = 95.392, p = .373, *partial* η^2 = .027.

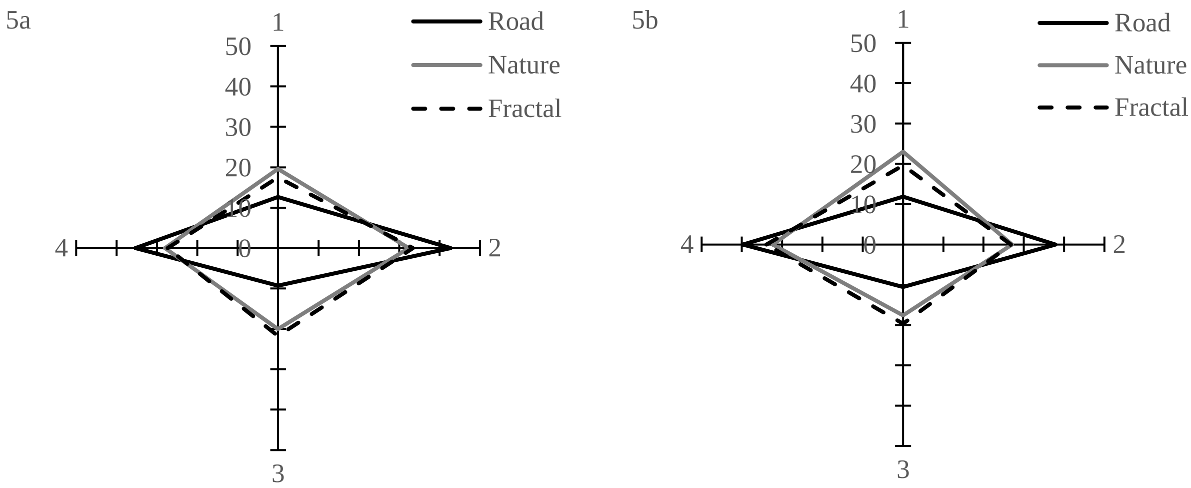

Figure 5: Percentage of saccades made in each of the four directions in (a) epoch 1 and (b) epoch 2 of the picture task, across the three conditions of image type. Bins 1, 2, 3, and 4 identified on the axis represent upwards, rightwards, downwards, and leftwards saccades respectively.

4. Discussion

Research has shown that top-down settings which guide attentional resources in one task can persist to a second task, influencing performance (e.g., Leber & Egeth, 2006; Longman et al., 2013; Thompson et al., 2007) and the spatial allocation of attention in the second task (e.g., Hills et al., 2016; Longman et al., 2017; Thompson & Crundall, 2011; Thompson et al., 2015; Wendt et al., 2017). The current study explored the underlying mechanisms associated with this “attentional inertia” effect. The work aimed to measure whether top-down contextual information would influence the persistence of a previously relevant set. An additional aim was to confirm the duration of carry-over of settings from a preceding task, replicating the work of Thompson et al. (2015). Participants searched through letter strings arranged horizontally, vertically, or randomly, and then made a complexity judgement about road images, nature scenes, or fractal images whilst their eye movements were measured. Based on past findings it was predicted that the orientation of the letter strings would affect eye movements in the pictures and that this effect would be brief. Crucially it was also hypothesised that carry-over would vary according to the image type in the second task.

The impact of top-down contextual influences present in the second task was investigated to explore the mechanisms associated with attentional set switching. It has been argued that switching an intentional (task) set incorporates both inhibition of a previously relevant task set and activation of a new task set (e.g. Rogers & Monsell, 1995). Studies measuring attentional inertia have found that persistence of attentional settings are greater with increased investment in the initial set (Leber & Egeth, 2006; Thompson et al., 2015) indicating that the ability to inhibit an attentional set affects switching (the more investment, the more difficult the set is to inhibit, therefore the greater the chance of carry-over). Additionally, attentional inertia is reduced when observers have time to prepare for a switch (Longman et al., 2017). Again, this would indicate that inhibition of previously relevant settings is an important component as the increased preparation time allows for more inhibition. However, research in this field has not yet considered the role of activation of the new set and it may be the case that, similar to an intentional set, the time taken to activate a new attentional set impacts on the ability to switch set.

It was reasoned that varying the contextual information present within the second task would impact the activation of an attentional set for this task and therefore have an influence on attentional inertia. Real-world images have “global scene properties” (Torralba et al.,

2006) that guide attention and eye movements. These properties (also referred to as the "scene gist") can be extracted very quickly, allowing observers to categorise natural scenes within 80ms of viewing (Wu, Wang, & Pomplun, 2014). Road and nature scenes tend to have a common structure, they are familiar, and studies show they evoke a stereotypical spread of attention and search (e.g., Crundall & Underwood, 1998; Foulsham et al., 2008; Shinoda et al., 2001). In contrast fractals lack semantic/contextual information (Foulsham & Kingstone, 2010). It was argued that the lack of context in fractal images would make it more difficult to extract scene gist and make it more difficult to activate an attentional set to guide resources. It was predicted that the carry-over of settings from the letter search to the picture search would therefore be greater in fractal images because activation of a suitable attentional set would take longer.

Despite this argument the present work found that semantic variations in the second task had no influence on carry-over from the letter search task. Patterns of eye movements across the three image types were markedly different with significantly wider horizontal search and significantly narrower vertical search on the road images compared to the nature and fractal images. This fits with past findings showing different eye movements across different image types (e.g., Crundall & Underwood, 1998; Foulsham & Kingston, 2010). Participants also rated the images differently in terms of complexity, judging the fractal images to be substantially more complex than the road or nature scenes. This shows that the different image types did have different influences upon the allocation of spatial attention. However, this did not interact with the carry-over of attentional set from the letter search task. Across all image types horizontal search was significantly wider following a horizontal letter search and significantly narrower following a vertical letter search. Whilst non-significant effects are difficult to interpret, calculation of Bayes factors provided very strong evidence for the null hypothesis. These results indicate that attentional inertia occurred throughout the experiment, despite varying the top-down contextual information in the picture search task. This would indicate that inhibiting previously relevant settings is more critical in the attentional inertia effect than activation of new settings.

It may however be argued that the manipulation of image type in the second task did not allow for a clear variation in contextual information. The data do show that participants rated the fractal images as more complex, however spread of search was similar for both nature scenes and fractal images (despite differences between fractals and road scenes). This does not necessarily support the proposal that fractals lack context, or that it takes longer to activate an attentional set with which to guide resources. Potentially the fact that the fractals were perceived to be more complex may have increased the top-down guidance of attention. In future a more effective manipulation may be to rotate or jumble the images, and the ease of categorising the different images could be measured to determine the speed with which observers can extract scene gist.

To fully investigate the impact of activation of a top-down set on attentional set switching it would also be important to manipulate the demands of the second task. Wu, Wick, and Pomplun (2014) state that top-down attentional guidance in real-world scenes can be separated into task demands and an understanding of the scene context. The present work measures how contextual information may bias attention and impact set switching, but this still leaves the role of specific task demands. Indeed, variations in the past findings provide some indication that task demand can influence set switching. The studies adopting the current experimental paradigm have used varying task instructions for the second task, viewing the images in preparation for a memory test (Thompson & Crundall, 2011), rating the images for hazardousness (Thompson et al., 2015), identifying hazards within road images (Hills et al., 2018), etc., and the way in which the carry-over effect manifests itself varies slightly across the different experiments. The effect is apparent in each of the previously published studies, yet in some experiments the letter search has more of an impact upon subsequent vertical spread of search and in some the orientation of the letters has a greater effect on the subsequent horizontal spread of search.

To illustrate this with an example, when asked to search road scenes for a memory test or to provide a rating of hazardousness the orientation of stimuli in a preceding letter task influenced the vertical spread of search but not the horizontal spread of search (Thompson & Crundall, 2011). In contrast, in the current study, when asked to rate a variety of images for complexity the orientation of preceding letters influenced both horizontal and vertical spread of search. Studies show that attentional resources are prioritised based on task demand (e.g., Hills et al., 2012; Shinoda et al., 2001) and the small differences in carry-over between the different experiments may imply that whilst the context of an environment may not directly affect the persistence of attentional settings, the task demands associated with that environment do interact with any previously relevant top-down settings. A criticism to this argument comes from the findings of Thompson et al. (2015) that showed both horizontal and vertical search in the picture task to be influenced by the orientation of the letters, despite the fact that participants were again searching road images to provide a rating of hazardousness. To expand on the present findings the impact of top-down influences in the second task should be more thoroughly investigated by varying both the stimuli and the search instructions in the second task.

A further limitation to the experimental paradigm is that the key press response options for both the letter search task and the picture search task are horizontally oriented. Findings show that when spatial response position is compatible to the stimulus position performance in a task can improve, even when the position of the stimulus has no relevance to the task. This is the stimulus response compatibility effect (e.g., Hommel & Prinz, 1997; Proctor & Vu, 2006). Whilst performance in the letter search and the ratings given to the pictures are not central to the investigation, it may be the case that the response options bias spatial attention along the horizontal axis. In the present study the carry-over effect was weaker in the measure of vertical spread of search, possibly indicating that the strength of the effect in the horizontal spread of search has been accentuated by the location of the response options. The previous results of Thompson et al. (2015) would argue against this however because using the same paradigm and a rating task in the picture search (with response options arranged horizontally) there was evidence of carry-over affecting both the horizontal and vertical spread of search. Nonetheless, this should be considered in future research to ensure that the attentional inertia effect can be studied independently of any impact of response mapping.

In accordance with past findings, the carry-over effect was small in this study and even those participants showing the greatest levels of attentional inertia were only making consistent eye movements across the letter and picture search 60% of the time. In addition, the costs of set switching were short-lived. Thompson et al. (2015) found significant carry-over of spatial settings within the first 1000ms of a second task but not in the second 1000ms and the present findings replicate this effect. The limited duration of the switch costs is also demonstrated in the study of Dombrowe, Donk, and Olivers (2011) who found that switching takes approximately 250-300ms. It is clear that carry-over of spatial attentional settings between two tasks has only a very small impact on the allocation of attentional resources, although the consistent finding of attentional inertia across multiple studies show that the effect should be considered when theorising about the influences on visual attention and search.

The current work shows that the top-down attentional settings used to complete one task can persist to a second task, causing a brief but noticeable effect on the spatial allocation of attention in this task. This supports past work in this field identifying attentional inertia as an independent influence on attention and performance. Moreover, the effect does not seem to be impacted by top-down contextual biases associated with stimuli in the second task. This would indicate that inhibition of a previously relevant set is more critical in attentional set switching than activation of the new set. This could be further explored by varying the demands of the second task in addition to the stimuli used in the second task.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421-452). Cambridge, MA: MIT.

Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition, 13,* 99–108.

Crundall, D. E., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics, 41,* 448-458.

Dombrowe, I., Donk, M., & Olivers, C. N. L. (2011). The costs of switching attentional sets. *Attention, Perception, & Psychophysics*, *73*(8), 2481-2488. doi: 10.3758/s13414-011-0198-3.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10 (1),* 32-39.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044.

Foulsham, T., & Kingstone, A. (2010). Asymmetries in the direction of saccades during perception of scenes and fractals: Effects of image type and image features. *Vision Research, 50,* 779–795.

Foulsham, T., Kingstone, A., & Underwood, G. (2008). Turning the world around: Patterns in saccade direction vary with picture orientation. *Vision Research, 48*(17), 1777-1790. doi: 10.1016/j.visres.2008.05.018.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108,* 316-355. doi: 10.1037/0096-3445.108.3.316.

Hills, P. J., Mileva, M., Thompson, C., & Pake, J. M. (2017). Carryover of scanning behaviour affects upright face recognition differently to inverted face recognition. *Visual Cognition.* doi: 10.1080/13506285.2017.1314399.

Hills, P. J., Sullivan, A. J., & Pake, J. M. (2012). Abberant first fixations when looking at inverted faces in various poses: The results of the centre-of-gravity effect? *British Journal of Psychology, 103,* 520–538.

Hills, P. J., Thompson, C., Jones, S. P., Piech, R. M., Painter, L., & Pake, J. M. (2016). Attentional modulation of the carry over of eye-movements between tasks. *Acta Psychologica, 167*, 1-15. doi: 10.1016/j.actpsy.2016.03.012.

Hills, P. J., Thompson, C., & Pake, J. M. (2018). Detrimental effects of carryover of eye movement behaviour on hazard perception accuracy: effects of driver experience, difficulty of task, and hazardousness of road. *Transportation Research Part F: Traffic Psychology and Behaviour, 58,* 906-916. doi: 10.1016/j.trf.2018.07.014.

Hommel, B., & Prinz, W. (1997). *Theoretical issues in stimulus-response compatibility.* Amsterdam, The Netherlands: North Holland.

JASP Team (2020). JASP (Version 0.13.1). Retrieved from https://jasp-stats.org/download/

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology, 37,* 43-75.

Kaptein, N. A., Theeuwes, J. & van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1053-1069.

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis and Prevention, 47*, 827–834.

Leber, A. B., & Egeth, H. E. (2006). Attention on autopilot: Past experience and attentional set. *Visual Cognition, 14,* 565–583.

Longman, C. S., Lavric, A., & Monsell, S. (2017). Self-paced preparation for a task switch eliminates attentional inertia but not the performance switch cost. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 43*(6), 862-873. doi: 10.1037/xlm0000347.

Longman, C. S., Lavric, A., & Monsell, S. (2013). More attention to attention? An eye-tracking investigation of perceptual attributes during a task switch. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1142–1151.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1-20.

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251-256). San Diego, CA: Elsevier.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic respresentation of the spatial envelope. *International Journal of Computer Vision, 42,* 145-175.

Proctor, R. W., & Vu, K-P. L. (2006). *Stimulus-response compatibility principles: Data, theory, and application.* Boca Raton, FL: CRC Press.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124 (2),* 207-231.

Rossi, A. F., & Paradiso, M. A. (1995). Feature-specific effects of selective visual attention. *Vision Research, 35,* 621-634.

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2002). Components of attentional set-switching. *Experimental Psychology, 52 (2),* 83-98.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. detection, search, and attention. *Psychological Review, 84 (1),* 1-66.

Shinoda, H., Hayhoe, M. M., & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research, 41,* 3535–3545.

Tatler, B. W., & Vincent, B. T. (2009). The prominence of behavioural biases in eye guidance. *Visual Cognition, 17*(6-7), 1029-1054. doi: 10.1080/13506280902764539.

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica, 83,* 93-154.

Thompson, C., & Crundall, D. (2011). Scanning behaviour in natural scenes is influenced by a preceding unrelated visual search task. *Perception, 40,* 1335-1349. doi: 10.1068/p6848.

Thompson, C., Howting, L., & Hills, P. (2015). The transference of visual search between two unrelated tasks: Measuring the temporal characteristics of carry-over. *Quarterly Journal of Experimental Psychology.* doi: 10.1080/17470218.2015.1013042.

Thompson, C., Underwood, G., & Crundall, D. (2007). Previous attentional set can induce an attentional blink with task-irrelevant initial targets. *Quarterly Journal of Experimental Psychology, 60*(12), 1603-1609. doi: 10.1080/17470210701536468.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766-786. doi: 10.1037/0033-295X.113.4.766.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problem of *p* values. *Psychonomic Bulletin and Review, 16*(2), 779-804. doi: 10.3758/BF03194105.

Wendt, M., Kähler, S. T., Luna-Rodriguez, A., & Jacobsen, T. (2017). Adoption of task-specific sets of visual attention. *Frontiers in Psychology, 8*, 687. doi: 10.3389/fpsyg.2017.00687.

Wu, C-C., Wang, H-C., & Pomplun, M. (2014). The roles of scene gist and spatial dependency among objects in the semantic guidance of attention in real-world scenes. *Vision Research, 105,* 10-20. doi: 10.1016/j.visres.2014.08.019.

Wu, C-C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology, 5,* 54. doi: 10.3389/fpsyg.2014.00054.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs”. *Psychological Research, 63,* 212-233.